

Stochastically Projecting Tensor Networks

Bryan K. Clark and Hitesh J. Changlani

Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green St, Urbana IL 61801, USA

(Dated: April 10, 2014)

We apply a series of projection techniques on top of tensor networks to compute energies of ground state wave functions with higher accuracy than tensor networks alone with minimal additional cost. We consider both matrix product states as well as tree tensor networks in this work. Building on top of these approaches, we apply fixed-node quantum Monte Carlo, Lanczos steps, and exact projection. We demonstrate these improvements for the triangular lattice Heisenberg model, where we capture up to 57% of the remaining energy not captured by the tensor network alone. We conclude by discussing further ways to improve our approach.

I. INTRODUCTION

Numerical methods are important tools for understanding strongly correlated systems. One such approach is the density matrix renormalization group (DMRG) [1, 2] which is particularly powerful for one-dimensional quantum systems. In recent years, DMRG has also had many successes in simulating quasi one-dimensional ladders which is a way of approaching the two-dimensional limit [3]. Unfortunately, the bond-dimension D required in DMRG (the parameter that determines the accuracy), scales exponentially with the width of the ladders. The underlying variational wavefunction that is optimized by the DMRG algorithm is a matrix product state (MPS) [4]. Higher dimensional generalizations of the MPS idea have led to development of new tensor network (or related) approaches [5] such as TTN [6–8], MERA [9], CPS [10–13] and PEPS [14, 15]. These are known to be better than DMRG at capturing the entanglement structure of physical systems and hence can produce better energies for the same bond-dimensions. Unfortunately, the computational expense of optimizing these wavefunctions and calculating observables (such as the energy) scales formidably with D . This motivates us to look for approaches which take a low bond-dimension ansatz generated by DMRG or other tensor network methods and improve upon them.

One such approach is to apply particular projectors \hat{P} to a tensor network wavefunction Ψ which generates a new wavefunction $\hat{P}\Psi$ which is closer to the ground state. In fact, this technique is often used to optimize a matrix product state through evolution in imaginary time [16]. In this paper, we will explore the ways in which quantum Monte Carlo methods can help generate or evaluate projectors \hat{P} producing a (possibly stochastic) representation of $\hat{P}\Psi$.

There has been fruitful progress when Monte Carlo and tensor network approaches have been combined. Examples include Refs. [17–19] where Monte Carlo is used to evaluate observables or optimize tensor networks and Ref. [20] where a MPS is used as one component of a larger variational ansatz. Additionally, for the square $J_1 - J_2$ model, Ref. [21] used an approach similar to one of the directions explored here.

In this paper, we explore the application of three projectors \hat{P} to tensor networks: a stochastic application of (1) $\hat{P}_{\text{exact}} = \exp[-\beta H]$ absorbing the sign problem for small imaginary time β , (2) $\hat{P}_{\text{Lanczos}} \equiv 1 + \alpha H + \Delta H^2$ [22, 23] (as

well as generalizations thereof) and (3) \hat{P}_{FN} , a sign problem free projector proposed and studied in Refs. [24, 25]. We also discuss the possibility of chaining these different projectors.

To demonstrate these approaches, we consider the spin 1/2 Heisenberg Hamiltonian

$$H = \sum_{\langle i,j \rangle} S_i \cdot S_j \quad (1)$$

on the two-dimensional triangular lattice where S_i refers to the spin 1/2 operator on site i and $\langle i, j \rangle$ are nearest neighbors. This system is frustrated and has a sign problem in the Ising basis. The triangular lattice has been commonly used as a benchmark for testing new tensor network algorithms [26, 27] and is considered challenging for these methods owing to the high coordination number of every site. We use cylindrical boundary conditions with an equal number of unit cells in each direction.

In the remainder of the paper, we will explain our methods and demonstrate that these projectors typically recover between 15 and 57 percent of the energy missed by the tensor networks themselves.

II. TENSOR NETWORK STATES

Consider the full many body wavefunction $|\Psi\rangle$ expanded in a basis of configurations $|c\rangle$ with coefficients $\Psi[c]$

$$|\Psi\rangle = \sum_c \Psi[c] |c\rangle \quad (2)$$

For a N spin system the chosen configurations could be Ising states $|c_1 c_2 \dots c_N\rangle$ with c_i referring to the spins on site i . Then a tensor network is a variational ansatz for $\Psi[c]$, which can be generically written as,

$$\Psi[c] = \prod_i M_{\vec{\alpha}}[i; c_i] \quad (3)$$

where each $M_{\vec{\alpha}}[i; c_i]$ is indexed by the site i and the value of the physical spin c_i . The values of $\vec{\alpha}$ are arbitrary tensor indices $\vec{\alpha} = \{\alpha_0 \dots \alpha_n\}$. In this notation, the product implies a contraction over all shared tensor indices between tensors. In any tensor network scheme, the indices are chosen so that contraction over them leaves $\Psi[c]$ a scalar. Each tensor index

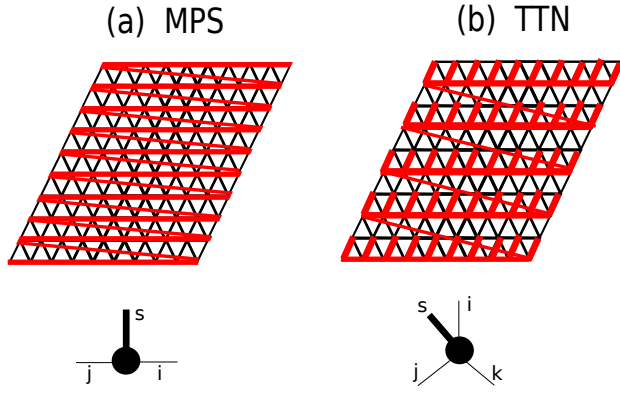


Figure 1: (Color online) Tensor networks shown in red used in this work (a) Matrix Product State which is optimized with the DMRG algorithm. This involves mapping the sites on a 2D lattice onto a "snake" used to define the sweep order. Each site has two matrices associated with it (one for each spin value) with auxiliary indices i and j which are summed over. (b) A particular realization of a tree tensor network with coordination 3 and coordination 1 vertices, which is a particular map of spins on a 2D lattice onto a tree (loop-less) network. Each coordination 3 site has three indices i, j, k which are summed over. The figure also shows the underlying triangular lattice (in black) on which the nearest neighbor Heisenberg Hamiltonian is defined. Periodic boundaries along the x-direction have not been shown.

has a maximal dimension of size D which is a tuning parameter that extrapolates between a product state ($D = 1$) and the exact wave-function ($D \rightarrow \infty$). Smaller D corresponds to lower entanglement. Different forms of tensor networks differ in the way in which the tensors are connected.

A. Matrix Product States (MPS) and Tree Tensor Network (TTN) States

In this work, we will focus on tensor networks whose indices are connected into paths (MPS) as well as those that are connected into trees (TTN). While MPS has been used extensively for a wide variety of 1D and quasi 2D systems [28], the use of TTN is relatively recent. Several authors have employed them to study tree geometries [29–32], 2D geometries and quantum chemistry [8, 33]. It is to be noted that the flavor of TTN used here (and in all the above mentioned works) is such that the physical spins occupy the vertices of a tree. [43]

Figure 1 shows the networks and the tensors we use. The trees we have selected have coordination 3 and 4. The MPS wavefunctions are generated using the ITensor library [34] and the TTN wavefunctions have been generated from our Tree-DMRG code, (details of which have already been discussed elsewhere [29]).

It is important to note that the "optimal mapping" from a tree to a 2D system is not a trivial problem, since there is a subtle interplay between short and long range entanglement necessary for getting gains in the energy. Practitioners of TTN have typically used heuristic algorithms to arrive at a reasonable mapping. Here we select a network that covers some

short range motifs not efficiently covered by the MPS. These networks can be contracted computationally efficiently using $O(D^z)$ operations with z being the maximum degree of the tree.

As an example, for $D = 8$, we find that switching from MPS to TTN can gain 30% (for coordination 3) and 40% (for coordination 4) of the missing energy.

While we focus on these particular tensor networks in this work, our conclusions likely generalize to other tensor networks such as PEPS, where contraction is even more computationally costly (exponential in D if performed exactly and D^6 if performed approximately [35]) and where it is even more important to improve on lower bond dimensions. Since in the limit of large D , any state can be represented by a tensor network Ψ , the state $\hat{P}\Psi$ is also guaranteed to converge to the correct energy in this limit.

B. Use of quantum numbers and gains from post projection

Since DMRG is a D^3 algorithm, several efforts are made in its implementation to reduce the prefactor associated with the calculation. One such improvement, at the cost of some extra book-keeping, is the use of the good quantum numbers such as total S_z and/or total particle number.

However, it is not apriori obvious that this is necessarily advantageous, because respecting a global symmetry may introduce additional entanglement in the state. Instead, breaking symmetries and then restoring them may allow for lowering of the energy for a given bond dimension. This has recently been observed in the context of Hartree Fock wavefunctions [36]. One can reinstate the symmetry using Quantum Monte Carlo (see Figure 3 for the gain in energy before and after projection), by sampling configurations which have definite S_z and/or particle number. We see evidence for this assertion, when we compare MPS with and without the use of any good quantum numbers (see Figure 3 and 4 for a comparison). Also note that such a global post projection of the tensor network is not necessarily very straightforward in the MPS framework alone, but easy to perform in QMC.

III. PROJECTOR QUANTUM MONTE CARLO

QMC methods work by filtering out excited states from a given trial state Ψ_T by applying to it, a projector,

$$\hat{P}|\Psi_T\rangle = \lim_{\beta \rightarrow \infty} \exp(-\beta H)|\Psi_T\rangle \propto |\Psi_0\rangle \quad (4)$$

where $|\Psi_0\rangle$ is the ground state. QMC methods are most efficient in systems where there is no "sign problem"; for example, when all the off-diagonal matrix elements of H are negative in a chosen basis. Otherwise, the stochasticity in the method causes the variance to grow exponentially with system size and inverse temperature β arising from a cancellation of large positive and negative terms. This makes it difficult to reliably measure observables for system sizes of interest.

However, as we will see later in this section, some approximate "sign-problem free" projectors can also be used, at the expense of a variational bias in the final answer.

To quantify the gains from a generic projector methods using a MPS wavefunction (henceforth abbreviated as QMC+MPS), we define the fraction of missing energy as

$$f = \frac{E_{MPS}[D] - E_{QMC+MPS}[D]}{E_{MPS}[D] - E_{MPS}[\infty]} \quad (5)$$

where $E_{MPS}[D]$ is the energy of the MPS of bond dimension D , $E_{QMC+MPS}[D]$ is the energy of the QMC using the same MPS of bond dimension D and $E_{MPS}[\infty]$ is taken to be the energy extrapolated to $D \rightarrow \infty$, which is found to be $-53.04 J$.

A. Exact Projection

Projector Monte Carlo methods work by producing a stochastic representation of the entire many body wavefunction. The important sampled wavefunction ($\Psi\Psi_T$) is represented by walkers, entities that carry weights w_c and signs s_c associated with configurations (or basis states, denoted by c) in the Hilbert space.

The projected (or mixed) energy can be written as

$$E = \frac{\langle \Psi | H | \Psi_T \rangle}{\langle \Psi | \Psi_T \rangle} \quad (6)$$

which can be recast as,

$$E = \frac{\sum_c |\Psi(c)\Psi_T(c)| \text{sgn}(\Psi(c)\Psi_T(c)) E_L[c]}{\sum_c |\Psi(c)\Psi_T(c)| \text{sgn}(\Psi(c)\Psi_T(c))} \quad (7)$$

where $|\Psi[c]\Psi_T[c]|$ is the probability distribution being sampled by the projection process and $E_L[c] \equiv \langle c | H | \Psi_T \rangle / \langle c | \Psi_T \rangle$ which is referred to as the "local energy". Thus, we compute the estimator

$$E = \frac{\langle s_i E_L[i] \rangle}{\langle s_i \rangle} \quad (8)$$

where i are the samples accumulated during the projection process. Since this estimator corresponds to the energy of the wavefunction $\exp(-\beta H/2)|\Psi_T\rangle$, it is guaranteed to be a variational upper bound of the true ground state. Notice that each calculation of the form $E_L[i]$ takes $O(N)$ contractions of Ψ_T where N is the number of sites.

We apply the exact projector (for small β) using (release-node) QMC [37] testing it on the 10×10 triangular lattice. This is accomplished by first sampling, via variational Monte Carlo, the distribution of walkers on configurations c from $|\langle \Psi_T | c \rangle|^2$. Our Ψ_T is the MPS wavefunction with $D = 16$ (without quantum numbers). The results are shown in Figure 2. The starting energy is the variational energy of this state, which is calculated to be around $-49.2 J$.

Once the exact projection is started, the energy is found to decrease but the statistical error quickly increases. We

find that beyond $\beta \approx 0.2$, the errorbars become unacceptably large. These runs have been calculated with 1.6×10^6 walkers. Error bars can be reduced with more cores (statistics) and efficient algorithms [38–40], but we do not expect to go to significantly larger β because the average sign decays to zero exponentially (as seen in Figure 2).

B. Fixed Node

Because exact QMC projection scales exponentially, applying \hat{P}_{exact} eventually becomes untenable. Instead, one can apply fixed node QMC (FNQMC) which takes a Hamiltonian H with a sign problem as well as a guess wavefunction Ψ_T for the ground state of H and generates a new (importance-sampled) effective Hamiltonian H_{FN} that has no sign problem [24, 25]. This new Hamiltonian has off diagonal elements,

$$(H_{FN})_{ij} = H_{ij} \frac{\psi_T(c_j)}{\psi_T(c_i)} \quad \text{if } H_{ij} \frac{\psi_T(c_i)}{\psi_T(c_j)} < 0 \quad (9a)$$

$$= 0 \quad \text{otherwise} \quad (9b)$$

and whose diagonal elements are modified as,

$$(H_{FN})_{ii} = H_{ii} + \sum_{j \in \text{s.v.}} H_{ij} \frac{\psi_T(c_j)}{\psi_T(c_i)} \quad (10)$$

where "s.v." refers to only the sign violating terms (terms that have $H_{ij} \frac{\psi_T(c_i)}{\psi_T(c_j)} > 0$).

The ground state Ψ_{FN} of H_{FN} is an approximation to the ground state Ψ_0 of H which becomes exact in the limit where $\Psi_T = \Psi_0$. In addition, the energy is variationally upper-bounded so that $\langle \Psi_{FN} | H_{FN} | \Psi_{FN} \rangle \geq \langle \Psi_{FN} | H | \Psi_{FN} \rangle \geq \langle \Psi_0 | H | \Psi_0 \rangle$. Therefore an effective use of FNQMC requires generating a good guess Ψ_T . The only requirement for Ψ_T is the ability to quickly evaluate the amplitude $\Psi_T[c]$ for arbitrary configurations c . The general lore is that the sign structure of Ψ_T is the most important aspect of the guess with a secondary importance on the amplitudes (in the continuum, it can be shown that only the sign structure matters). We use tensor networks, then, for Ψ_T .

As shown in Figure 4, we apply FN-QMC on MPS and TTN of fixed quantum numbers ($S_z = 0$) with various bond dimensions. Calculating $\langle \Psi_{FN} | H_{FN} | \Psi_{FN} \rangle$, we find that the tensor network ansatz for low bond dimensions lowers the energy by about $1J$. The tree ansatz provides a better starting point for low D and improves upon increasing coordination. These ansatz are able to capture about 15-20% of the remaining missing energy.

We find that the quantum number free MPS (followed by QMC post projection), provides a much better starting point than its quantum number counterpart, and with FNQMC captures a much larger fraction of the energy, about 40%.

Note that the fixed node energies underestimate the quality of the wavefunction as the true energy, $\langle \Psi_{FN} | H | \Psi_{FN} \rangle$ is actually even closer to the ground state energy.

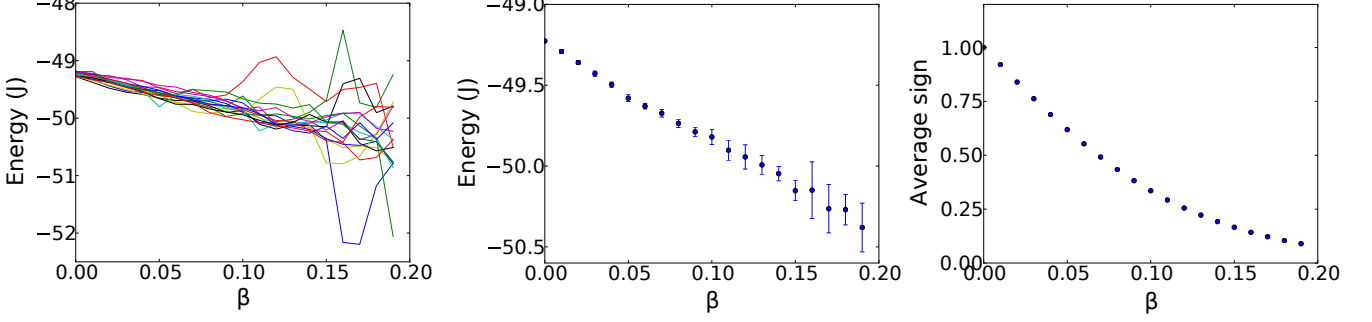


Figure 2: Left: Energy of $\exp[-\beta H/2]|\Psi_T\rangle$ as a function of β computed using projector quantum Monte Carlo for 16 different runs of 100,000 walkers each. Center: The energy vs β after averaging the runs. Right: Average sign $\langle s_i \rangle$ as a function of β computed using projector quantum Monte Carlo. The average sign decays exponentially. The trial wavefunction used in the initial VMC is the $D = 16$ MPS with no quantum numbers.

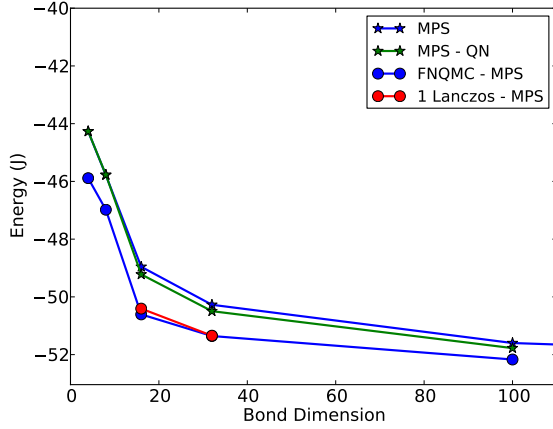


Figure 3: Total ground state energy obtained from MPS of various bond dimensions generated without the use of total S_z quantum number. The projection onto total $S_z = 0$ sector is reinstated in the QMC calculation by calculating the energy over samples with that symmetry. The legend shows different projection methods and their respective decrease in energy as a function of bond-dimension.

C. Lanczos Step

The final projector we consider is applying \hat{P}_{Lanczos} . In order to apply this, we need to compute the optimal values of α, Δ which we take to be those that minimize the variational energy. Consider the energy

$$\langle E \rangle = \frac{\langle \Psi(1 + \alpha H + \Delta H^2) | H | (1 + \alpha H + \Delta H^2) \Psi \rangle}{\langle \Psi(1 + \alpha H + \Delta H^2) | (1 + \alpha H + \Delta H^2) \Psi \rangle} \quad (11)$$

This requires computing terms of the form $\langle \Psi | H^n | \Psi \rangle$ for $n \leq 5$. Computing these terms can be done either in the context of a matrix-product state/matrix-product operator formalism or within quantum Monte Carlo. In the MPS/MPO formalism, one can generate a MPO operator for H which consists of four matrices M^{σ_i, σ_j} for each site, each of bond dimension χ (for the 10×10 lattice, $\chi = 35$). One can then

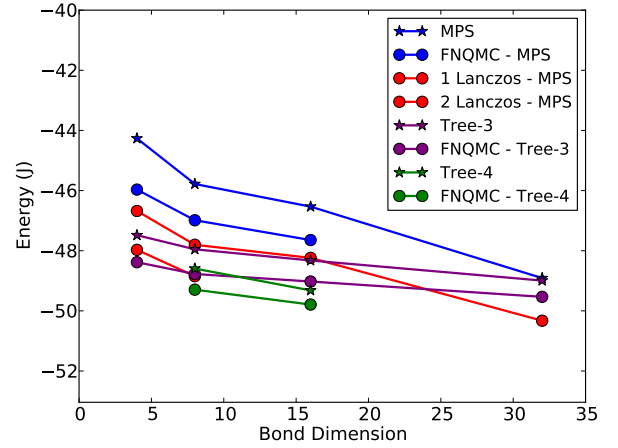


Figure 4: Total ground state energy obtained from MPS and TTN of various bond dimensions generated with the use of total S_z quantum number. Legend shows different projection methods and their respective decrease in energy as a function of bond-dimension.

apply the MPO to the MPS computing a new MPS $H^{n/2}|\Psi\rangle$ by contracting over the physical degrees of freedom. This new MPS will have a bond dimension of $\chi^n/2D$. Then to compute $\langle \Psi H^n \Psi \rangle$ we simply compute the overlap of $|H^{n/2}\Psi\rangle$ with itself (appropriate modifications can be made for odd n). Using this viewpoint, the computational complexity of this procedure is then $(D^3 \chi^{3n/2})$ but the constant prefactor for this approach is extremely good because it requires only a single overlap calculation. Further improvements from the sparsity of the MPO can potentially speed this up. Notice that this is significantly cheaper than doing a DMRG calculation at that bond-dimension.

One can alternately compute these terms using QMC. With variational Monte Carlo, one can compute $\langle \Psi_T H^n \Psi_T \rangle$ by sampling configurations R with probability $|\langle R | \Psi_T \rangle|^2$ and computing $\langle H^n | \Psi_T \rangle / |\Psi_T\rangle_{|\Psi_T|^2}$. The asymptotic scaling of these operations is $(6N)^n$ times the cost of contracting the tensor network where N is the number of sites. Hybrid ap-

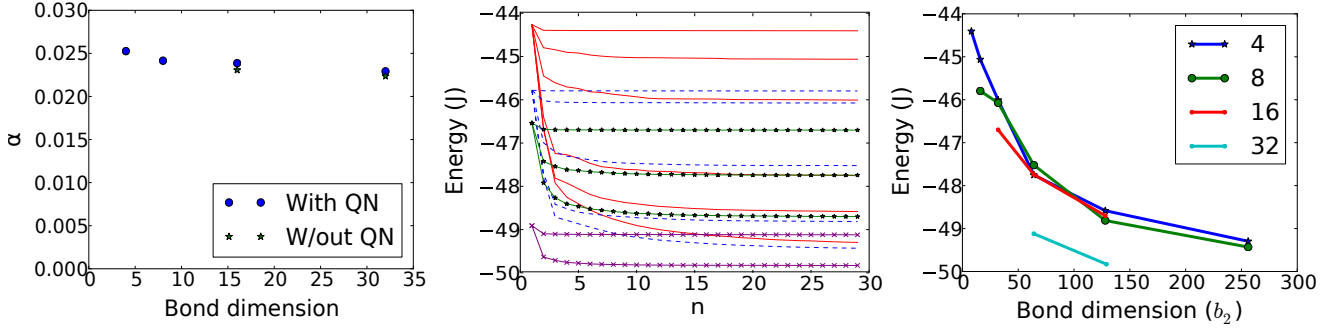


Figure 5: Left: Optimal exact 1-Lanczos wavefunction parameter α obtained by minimizing the variational energy, as a function of the bond dimension b_1 of the MPS. Center: Energy as a function of basis elements n in the approximate Lanczos method. The various colors correspond to choices of $b_1 = \{8 \text{ (solid red), } 16 \text{ (dotted blue), } 32 \text{ (starred green), } 64 \text{ (purple x's)}\}$. For each color/symbol the lines from top to bottom represent choice of b_2 starting from $2b_1$ and incrementing in multiples of 2. Right: Energy for $n = 30$ basis elements starting with an MPS of bond dimension b_1 (shown in legend) as a function of bond dimension b_2 (x-axis).

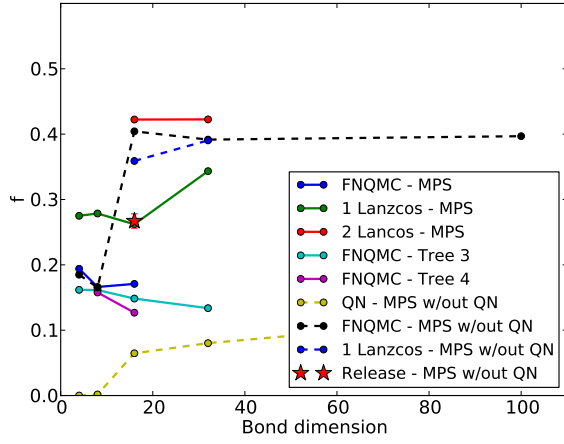


Figure 6: Fraction of the energy f (see eqn. (5)) recovered using various projection approaches on MPS and TTN as a function of bond dimension D .

proaches using some QMC and some MPO/MPS formalism are possible. The Lanczos numbers reported here were computed partially in the MPO-MPS formalism and partially with a hybrid MPO-QMC approach.

As Figure 6 shows (with use of quantum numbers), the 1-Lanczos step results are better than the fixed node results (albeit more expensive), recovering approximately 30% of the missing energy. The 2-Lanczos does even better recovering 40%. Interestingly, without quantum numbers, the Lanczos step does approximately as well as fixed-node but this underestimates the quality because our present implementation of this method does not perform a post projection onto the definite S_z sector.

Finally, we note that because of the computational cost, it is difficult to go beyond a few Lanczos steps. This is essentially because we are finding optimal parameters in the exact basis of $H^n|\Psi\rangle$. A possible alternative is to generate a basis where each basis element is a cheap approximation to $H^n|\Psi\rangle$. One

way to accomplish this is to repeatedly apply the MPO operator H to $|\Psi\rangle$ of bond dimension b_1 truncating down to a fixed bond-dimension b_2 at each step using the zip-up method [41] (see fig. 5 (center and right)). The Hamiltonian and overlap matrices can then be generated in this basis and a variational upper bound to the ground state can be found by solving a generalized eigenvalue problem. This can do no better than DMRG on a starting bond dimension of nb_2 . Nonetheless, n can in principle be quite large since one simply needs to compute the n^2 matrix elements. Empirically, though, we find that using this approach, we do not gain significant energies for $n > 4$. This is confirmed by looking at the (numerical) rank of the system. Interestingly enough, we find that, at least until $b_1 > 16$, the value of b_1 doesn't matter significantly. We find a better energy using approximate Lanczos for $b_1 = 4$ and $b_2 = 64$ compared to exact 1-Lanczos on $b_1 = 4$ which uses a MPS $H|\Psi\rangle$ of bond dimension 140. To compete with the 2-Lanczos on $b_1 = 4$ which uses a MPS of bond dimension 4900, we need only $b_2 = 128$. At $b_1 = 4, b_2 = 256$ we actually restore 57% of the missing energy in the tensor network.

IV. CONCLUSION

In conclusion, we have shown how to improve upon tensor networks using projection techniques in a way that is efficient and massively parallelizable (the QMC methods are all largely embarrassingly parallel). We have tested stochastic exact projection, Fixed Node QMC, and Lanczos steps done in different ways using tensor networks. We summarize our findings using the 10×10 triangular lattice (with cylindrical boundaries) nearest neighbor Heisenberg model as a test bed.

We have found that it is feasible to perform release node calculations starting with MPS of reasonably small bond dimensions. We released the walk after starting with a VMC calculation and gained about 20% of the missing energy, before the numerical sign problem took over. We expect to achieve further improvements by chaining projectors. For example,

one can start with a distribution drawn from a fixed-node or Lanczos calculation.

Next, we applied a sign problem free fixed-node QMC projector (for lattice systems) which is formulated to provide a variational upper bound to the energy. For MPS wavefunctions, this was found to be equivalent to effectively increasing D by a factor of 2 to 3. In order to provide a better starting wavefunction for low bond dimensions, we also used tree tensor networks of coordination 3 and 4. We believe that the gains reported in this paper can further be improved with optimized mappings of a 2D system onto a tree [33]. From a qualitative viewpoint, our results indicate that applying the fixed-node method to tensor networks such as PEPS (which have the correct entanglement laws and respect the symmetries of the lattice, but which scale unfavorably with D) should help in increasing the effective bond dimension. We also note the gains we achieved by post-projecting the quantum number free MPS onto the correct S_z sector, which is carried out in QMC in a straightforward way.

A third direction explored in this paper is the 1 (and 2) Lanczos step methods, which work by finding an optimal wavefunction in the basis spanned by $H\Psi$ (and $H^2\Psi$). Higher powers of $H^n\Psi$ were also calculated approximately in a MPO framework. While the 1-Lanczos was found to be close to the

fixed-node improvement, the 2-Lanczos appears to be generically better. In addition, we find significant improvement using the approximate Lanczos technique.

Added note: During the completion of this work, a paper by Wouters et al. [42] was posted to the arxiv exploring yet a fourth method to combine QMC with projection, in this case the AFQMC method. They find similar levels of improvements on a different model and slightly smaller system sizes.

V. ACKNOWLEDGEMENT

This research is part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (award number ACI 1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications. Computation was also done on Taub (UIUC NCSA). We acknowledge support from grant DOE, SciDAC FG02-12ER46875. We thank Cyrus Umrigar for discussions and critically reading the manuscript. We thank Michael Kolodrubetz and Katie Hyatt for useful conversations and Miles Stoudenmire for help with ITensor.

-
- [1] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
 - [2] S. R. White, Phys. Rev. B **48**, 10345 (1993).
 - [3] E. Stoudenmire and S. R. White, Annual Review of Condensed Matter Physics **3**, 111 (2012).
 - [4] S. Östlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995).
 - [5] F. Verstraete, V. Murg, and J. Cirac, Advances in Physics **57**, 143 (2008).
 - [6] Y.-Y. Shi, L.-M. Duan, and G. Vidal, Phys. Rev. A **74**, 022320 (2006).
 - [7] L. Tagliacozzo, G. Evenbly, and G. Vidal, Phys. Rev. B **80**, 235127 (2009).
 - [8] V. Murg, F. Verstraete, O. Legeza, and R. M. Noack, Phys. Rev. B **82**, 205105 (2010).
 - [9] G. Vidal, Phys. Rev. Lett. **101**, 110501 (2008).
 - [10] D. A. Huse and V. Elser, Phys. Rev. Lett. **60**, 2531 (1988).
 - [11] H. J. Changlani, J. M. Kinder, C. J. Umrigar, and G. K.-L. Chan, Phys. Rev. B **80**, 245116 (2009).
 - [12] F. Mezzacapo, N. Schuch, M. Boninsegni, and J. I. Cirac, New J. Phys. **11**, 083026 (2009).
 - [13] K. H. Marti *et al.*, New Journal of Physics **12**, 103008 (2010).
 - [14] A. Gendiar, N. Maeshima, and T. Nishino, Prog. Theor. Phys. **110**, 691 (2003).
 - [15] T. Nishino and K. Okunishi, Journal of the Physical Society of Japan **67**, 3066 (1998).
 - [16] G. Vidal, Phys. Rev. Lett. **93**, 040502 (2004).
 - [17] L. Wang, I. Pižorn, and F. Verstraete, Phys. Rev. B **83**, 134421 (2011).
 - [18] A. J. Ferris and G. Vidal, Phys. Rev. B **85**, 165147 (2012).
 - [19] A. W. Sandvik and G. Vidal, Phys. Rev. Lett. **99**, 220602 (2007).
 - [20] C.-P. Chou, F. Pollmann, and T.-K. Lee, Phys. Rev. B **86**, 041105 (2012).
 - [21] M. d. C. de Jongh, J. Van Leeuwen, and W. Van Saarloos, Phys. ical Review B **62**, 14844 (2000).
 - [22] E. Heeb and T. Rice, Zeitschrift für Physik B Condensed Matter **90**, 73 (1993).
 - [23] S. Sorella, Phys. Rev. B **64**, 024512 (2001).
 - [24] H. J. M. van Bommel *et al.*, Phys. Rev. Lett. **72**, 2442 (1994).
 - [25] D. F. B. ten Haaf *et al.*, Phys. Rev. B **51**, 13039 (1995).
 - [26] E. M. Stoudenmire and S. R. White, Phys. Rev. B **87**, 155137 (2013).
 - [27] S. R. White and A. L. Chernyshev, Phys. Rev. Lett. **99**, 127004 (2007).
 - [28] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005).
 - [29] H. J. Changlani, S. Ghosh, C. L. Henley, and A. M. Läuchli, Phys. Rev. B **87**, 085107 (2013).
 - [30] H. J. Changlani, S. Ghosh, S. Pujari, and C. L. Henley, Phys. Rev. Lett. **111**, 157201 (2013).
 - [31] S. Depenbrock and F. Pollmann, Phys. Rev. B **88**, 035138 (2013).
 - [32] W. Li, J. von Delft, and T. Xiang, Phys. Rev. B **86**, 195137 (2012).
 - [33] N. Nakatani and G. K.-L. Chan, The Journal of Chemical Physics **138**, (2013).
 - [34] The ITensor library is a freely available code developed and maintained on <http://itensor.org/index.html>.
 - [35] L. Wang, I. Pižorn, and F. Verstraete, Phys. Rev. B **83**, 134421 (2011).
 - [36] G. E. Scuseria *et al.*, The Journal of Chemical Physics **135**, (2011).
 - [37] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
 - [38] G. H. Booth, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009).
 - [39] F. R. Petruziello *et al.*, Phys. Rev. Lett. **109**, 230201 (2012).
 - [40] M. Kolodrubetz and B. K. Clark, Phys. Rev. B **86**, 075109 (2012).

- [41] E. Stoudenmire and S. R. White, New Journal of Physics **12**, 055026 (2010).
- [42] S. Wouters, B. Verstichel, D. Van Neck, G. K-L Chan, arXiv:1403.3125.
- [43] The flavor of TTN used here is quite distinct from the TTN

introduced by Vidal [7], which has all the physical sites as the leaves of a tree and a hierarchical arrangement of additional tensors that introduce entanglement. This is basically a MERA network [9] without the disentangling operations.